AI4Debunk
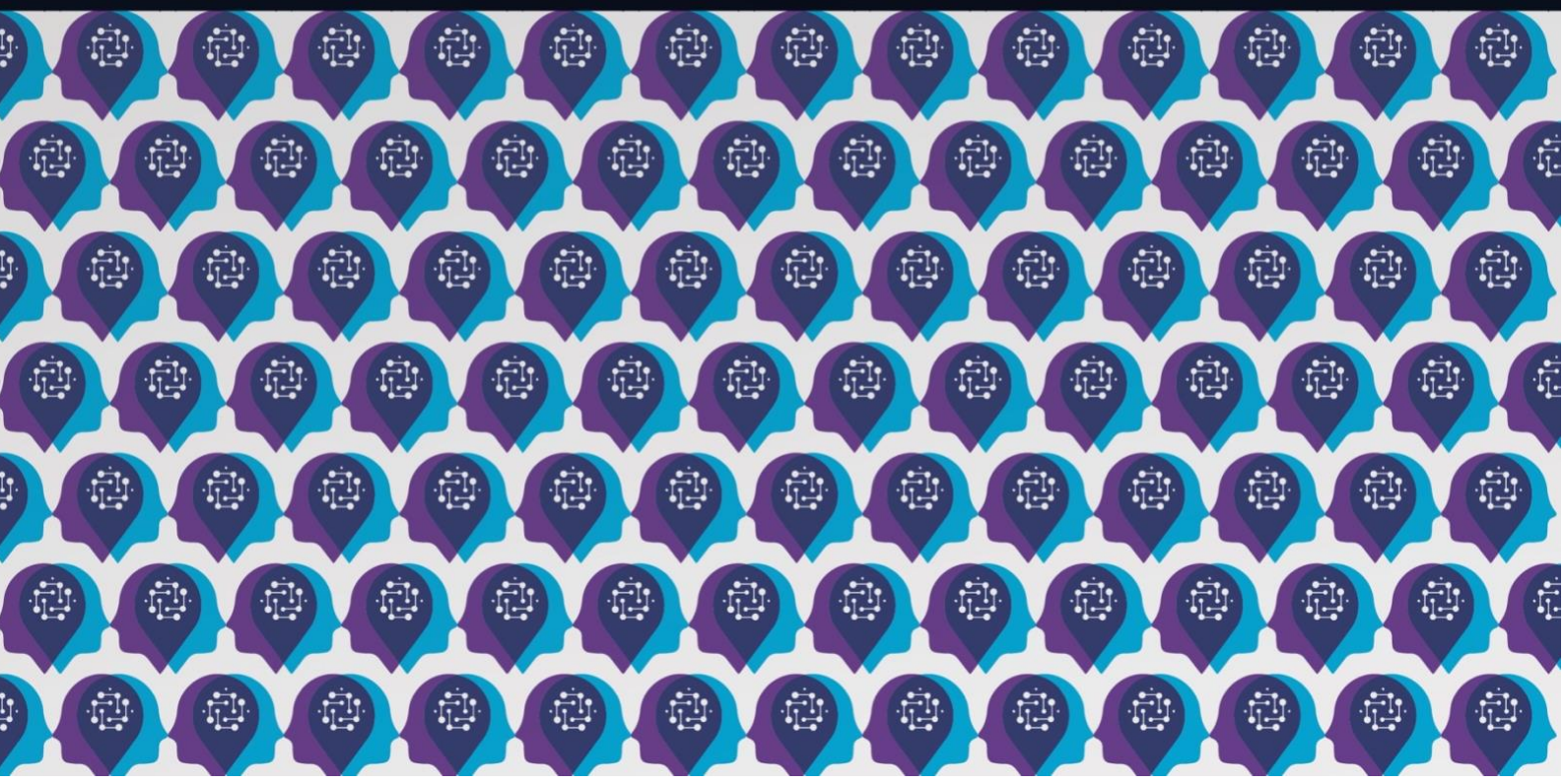
D6.1 STARTING DATASET OF FAKE STATEMENTS AND RELATED MULTIMEDIA CONTENTS

January 2025 (3rd release)

Grant Agreement No.: 101135757
Call: HORIZON-CL4-2023-HUMAN-01-CNECT
Topic: HORIZON-CL4-2023-HUMAN-01-05
Type of action: HORIZON Innovation Actions

D6.1 STARTING DATASET OF FAKE STATEMENTS AND RELATED MULTIMEDIA CONTENTS

final release

Project Acronym	AI4Debunk
Project Number	101135757
Project Full Title	Participative Assistive AI-powered Tools for Supporting Trustworthy Online Activity of Citizens and Debunking Disinformation
Work package	WP 6
Task	Task 1
Due date	31/01/2025
Submission date	31/01/2025
Deliverable lead	Partner CNR
Version	v1.0
Authors	Arianna D'Ulizia (Partner CNR), Alessia D'Andrea (Partner CNR), Marco Pirrone (Partner CNR)
Contributors	Alona Hryshko (Internews Ukraine), Georgi Gotev (Euractiv Bulgaria)
Reviewers	Pascaline Gaborit (Partner Pilot4dev)
Abstract	The deliverable D6.1 - Starting dataset of fake statements and related multimedia contents – describes the datasets prepared for collecting fake statements and related multimedia contents on the two case studies considered within the project (i.e., the war in Ukraine and the climate changes). A brief introduction on the objectives and expected outcome of Task 6.1 - Preparation of the starting dataset of fake statements and related multimedia contents – is

provided. Moreover, the structure and implementation of the two datasets are discussed, focusing on (i) the template used for extracting data from disinformation cases, (ii) the Database Management System (DBMS) used for the implementation of the datasets, and (iii) the cloud environment used for storing and sharing the multimedia contents. Afterward, a brief description of disinformation cases collected and analysed for the three releases of the datasets is provided. Finally, an assessment of the data quality of the datasets is carried out.

Keywords	Dataset, fake statements, multimedia content, disinformation cases

DOCUMENT DISSEMINATION LEVEL

Dissemination level	
X	PU - Public
	SEN - Sensitive

DOCUMENT REVISION HISTORY

Version	**Date**	**Description of change**	**List of contributor(s)**
0.1	16/12/2024	First draft	CNR, IUA, Euractiv Bulgaria
0.2	23/12/2024	Draft revised by the Task 6.1 participants	CNR, IUA, Euractiv Bulgaria, BSC, MICC/UNIFI, CNIT
0.3	17/01/2025	Final draft version revised by the internal reviewer	CNR, P4D
1.0	24/01/2025	Final version	CNR

STATEMENT ON MAINSTREAMING GENDER

The AI4Debunk consortium is committed to including gender and intersectionality as a transversal aspect in the project's activities. In line with EU guidelines and objectives, all partners – including the authors of this deliverable – recognise the importance of advancing gender analysis and sex-disaggregated data collection in the development of scientific research. Therefore, we commit to paying particular attention to including, monitoring, and periodically

**Funded by
the European Union**

evaluating the participation of different genders in all activities developed within the project, including workshops, webinars and events but also surveys, interviews and research, in general. While applying a non-binary approach to data collection and promoting the participation of all genders in the activities, the partners will periodically reflect and inform about the limitations of their approach. Through an iterative learning process, they commit to plan and implement strategies that maximise the inclusion of more and more intersectional perspectives in their activities.

DISCLAIMER

The AI4Debunk project has received funding from the European Union's Horizon Europe Programme under the Grant Agreement No. 101135757.

Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

COPYRIGHT NOTICE

How to cite this report: Arianna D'Ulizia, Alessia D'Andrea, Marco Pirrone, Alona Hryshko, Georgi Gotev (2024). AI4Debunk D6.1: Starting dataset of fake statements and related multimedia contents.with. *Link from website when deliverable is public*.

The AI4Debunk consortium is the following:

Participant number	Participant organisation name	Short name	Country
1	LATVIJAS UNIVERSITATE	UL	LV
2	FREE MEDIA BULGARIA	EURACTIV	BE
3	PILOT4DEV	P4D	BE
4	INTERNEWS UKRAINE	IUA	UA
5	CONSIGLIO NAZIONALE DELLE RICERCHE	CNR-IRPPS	IT
6	UNIVERSITA DEGLI STUDI DI FIRENZE	MICC/UNIFI	IT
6.1	CONSORZIO NAZIONALE INTERUNIVERSITARIO PER LE TELECOMUNICAZIONI	CNIT	IT
7	BARCELONA SUPERCOMPUTING CENTER CENTRO NACIONAL DE SUPERCOMPUTACION	BSC	ES
8	DOTSOFT OLOKLIROMENES EFARMOGES DIADIKTIOY KAI VASEON DEDOMENON AE	DOTSOFT	EL
9	UNIVERSITE DE MONS	UMONS	BE
10	NATIONAL UNIVERSITY OF IRELAND GALWAY	NUIG	IE
11	STICHTING HOGESCHOOL UTRECHT	HU	NL
12	STICHTING INNOVATIVE POWER	IP	NL
13	F6S NETWORK IRELAND LIMITED	F6S	IE

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

WP	Work Package
EC	European Commission
DBMS	Database Management System
EDMO	European Digital Media Observatory
KGR	Knowledge Graph Reasoning
DQAF	Data Quality Assessment Framework

EXECUTIVE SUMMARY

The deliverable D6.1 - Starting dataset of fake statements and related multimedia contents – describes the datasets prepared for collecting fake statements and related multimedia contents on the two case studies considered within the project (i.e., the war in Ukraine and the climate changes).

A brief introduction on the objectives and expected outcome of Task 6.1 - Preparation of the starting dataset of fake statements and related multimedia contents – is provided. Moreover, the structure and implementation of the two datasets are discussed, focusing on (i) the template used for extracting data from disinformation cases, (ii) the Database Management System (DBMS) used for the implementation of the datasets, and (iii) the cloud environment used for storing and sharing the multimedia contents. Afterward, a brief description of disinformation cases collected and analysed for the three releases of the datasets is provided. Finally, an assessment of the data quality of the datasets is carried out.

1   INTRODUCTION

This deliverable introduces the objectives and expected outcomes of Task 6.1 - Preparation of the starting dataset of fake statements and related multimedia contents - within Work Package 6 - Design, creation, and adaptation of knowledge graphs. Two different datasets have been prepared collecting fake statements and related multimedia contents on the case studies considered within the project (i.e., the war in Ukraine and the climate changes). Three releases of the datasets are expected in the project at M6, M9, and M13.

In this deliverable, the structure and implementation of the two datasets are described. In particular, the template used for extracting data from disinformation cases, the Database Management System (DBMS) used for the implementation of the datasets, as well as the cloud environment used for storing and sharing the multimedia contents are illustrated. A brief description of disinformation cases collected and analysed for the three releases of the datasets is provided. Finally, an assessment of the data quality of the datasets is carried out.

1.1   OBJECTIVES

This deliverable aims to describe the structure and implementation of two datasets (one on the war in Ukraine and the other on climate changes) containing a starting set of fake statements and multimedia contents (videos, images, audios, etc.), extracted both (i) after monitoring of websites and social media platforms that disseminate disinformation and fake news; and (ii) from a number of highly-reputable fact-checking websites (e.g., EDMO, Skepticalscience.com, Science.feedback.org) and databases owned by EURACTIV Bulgaria and Internews Ukraine.
From these gathered data, relevant information has been also extracted, such as the textual statement of disinformation, the author, the source, the date of publication, the topic, the keywords, the language, the fact-checking analysis, the rating scale, etc.

In Task 6.2, the information collected in the two datasets will be enriched with further relevant features (e.g. sentiment, higher-level features from eye-tracking, face recognition, voice analysis, etc.) extracted using machine learning and artificial intelligence tools.
In Task 6.3, the enriched datasets will be used as a starting knowledge base for the construction of the knowledge graphs that illustrate the structure of the deceptive data.

1.2   EXPECTED OUTCOME

Three releases of the datasets are expected in the project. The first release is planned at M6 (June 30, 2024), the second release is planned at M9 (September 30, 2024), and the third release is planned at M13 (January 31, 2025). All the information collected in the datasets is meant to be suitable for data processing and machine learning (ML) training and for the development of the semantic Knowledge Graph during the next steps of the AI4Debunk project.

2   DATASET STRUCTURE AND IMPLEMENTATION

In this section, the structure and implementation of the two datasets are described. In particular, the template used for extracting data from disinformation cases, the Database Management System (DBMS) used for the implementation of the datasets, as well as the cloud environment used for storing and sharing the multimedia contents are illustrated.
The first release of the datasets is available at the following link:

https://universityoflatvia387.sharepoint.com/:f:/r/sites/UG_AL4DEBUNK/Shared%20Documents/General/Work%20Packages/WP6%20-%20Design,%20creation,%20and%20adaptation%20of%20knowledge%20graphs/Deliverables/Deliverable%206.1%20-%20third%20release%20(January%202025)/DATASET%20final%20release%20(January%202025)?csf=1&web=1&e=kOBooT

The second release of the datasets is available at the following link:

https://universityoflatvia387.sharepoint.com/:f:/r/sites/UG_AL4DEBUNK/Shared%20Documents/General/Work%20Packages/WP6%20-%20Design,%20creation,%20and%20adaptation%20of%20knowledge%20graphs/Deliverables/Deliverable%206.1%20-%20second%20release%20(Sep%2724)/DATASET%20release%200.2%20(September%202024)?csf=1&web=1&e=whmMcj

The third release of the datasets is available at the following link:

https://universityoflatvia387.sharepoint.com/:f:/r/sites/UG_AL4DEBUNK/Shared%20Documents/General/Deliverables%20%26%20Milestones/2024/D6.1_Starting%20dataset/Deliverable%206.1%20-%20Release%200.3%20(January%202025)/DATASET%20final%20release%20(January%202025)?csf=1&web=1&e=Q7wXOp

2.1   TEMPLATE FOR INFORMATION EXTRACTION

The information extraction process was designed around some specific descriptive criteria; in particular, several parts of the fake news analysed have been selected to convert them into dataset fields. To achieve this task, a specific template was created and approved by Work Package (WP) 6.1.

In particular, the template was designed to extract the following information from each single claim collected from fact-checking sites, as shown in Table 1.

Table 1. TEMPLATE USED FOR INFORMATION EXTRACTION FROM DISINFORMATION CASES

FIELD	DESCRIPTION	VALUE
(TOPIC)	TOPIC OF THE DISINFORMATION (WAR IN UKRAINE OR CLIMATE CHANGES)	TEXT
(KEYWORDS)	KEYWORDS OF THE DISINFORMATION	TEXT
(DATE)	DATE OF PUBLICATION OF THE DISINFORMATION	DATE
(SOURCE)	MEDIA/PLATFORM/WEBSITE REPORTING THE DISINFORMATION	TEXT
(URL)	LINK TO THE DISINFORMATION	URL
(LANGUAGE)	LANGUAGE OF THE DISINFORMATION	TEXT
(AUTHOR)	AUTHOR OF THE DISINFORMATION	TEXT
(RATING SCALE)	RATING SCALE USED TO ASSESS TRUTHFULNESS	TEXT
(TEXT)	TEXTUAL STATEMENT OF THE DISINFORMATION IN ENGLISH	TEXT
(WHY)	FACT-CHECKING ANALYSIS IN ENGLISH	TEXT
(TEXT: SOURCE LANGUAGE)	TEXTUAL STATEMENT IN THE SOURCE LANGUAGE	TEXT
(WHY: SOURCE LANGUAGE)	FACT-CHECKING ANALYSIS IN THE SOURCE LANGUAGE	TEXT
(MULTIMEDIA)	AUDIOS, VIDEOS, IMAGES RELATED TO THE DISINFORMATION	URL

Source: Created by the authors

To identify the relevant fields that characterize the template, the fake news characterization introduced by Zhang & Ghorbani (2019) and refined by D'Ulizia et al. (2021) have been considered. In these works, the authors introduced a clear characterization of online fake news by identifying relevant features related to the users, content, and context that can be adapted to characterize also the datasets for the AI4Debunk project.

The field "multimedia" contains the URL to a folder available on the cloud storage services *Mediafire* (https://www.mediafire.com) or *Google Drive,* where the multimedia files (png, jpeg, mp4, etc.) are stored.

All the information extracted is meant to be suitable for data processing and machine learning (ML) training during the next steps of the AI4Debunk project. In particular, the fields (TOPIC), (KEYWORDS), (DATE), (SOURCE), (URL), (LANGUAGE), and (AUTHOR) have been collected for a proper assessment of the disinformation, while the remaining fields, such as (RATING SCALE), (TEXT), (WHY) and (MULTIMEDIA) would contribute to train ML algorithms and to develop a semantic Knowledge Graph. Note that the rating scale is assigned to the claims according to two different methods: 1) by extracting fake claims from certified and recognized fact-checking websites (please, see sub-sections 3.1.1, 3.2.1, 4.1.1 and 4.2.1); 2) by debunking fake claims with evidence and statements from reliable sources by expert journalists (please, see sub-sections 3.1.2, 3.1.3, 3.2.2, 4.1.2, 4.1.3, and 4.2.2).

The Knowledge Graph Reasoning (KGR) would be capable of deducing, assessing, and classifying new information based on the data present in the existing semantic base. Furthermore, the fields (TEXT: SOURCE LANGUAGE) and (WHY: SOURCE LANGUAGE) have been added to provide a multilanguage support to the Knowledge Graph.

Fig. 1 shows a sample screenshot related to a disinformation case reported on a fact-checking website, while in Fig. 2 an example of information extraction from a disinformation case is reported.
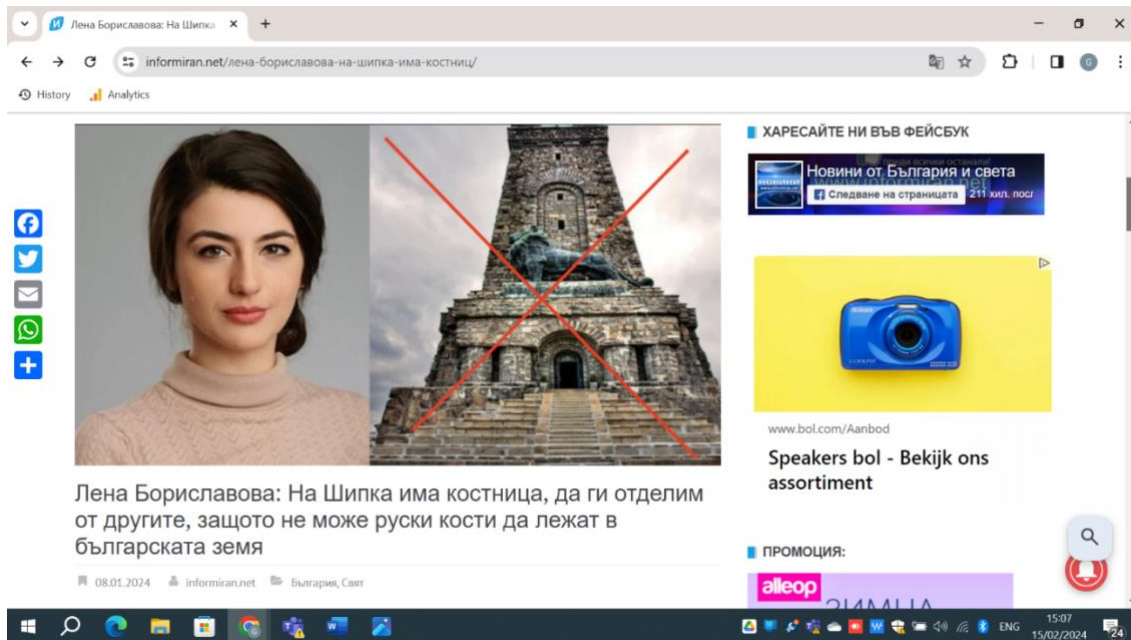
FIG. 1 EXAMPLE OF DISINFORMATION CASE REPORTED IN A FACT-CHECKING WEBSITE

8 January 2024

The author of the disinformation (AUTHOR):

The article is signed Desislava Stancheva (possibly a pen name)

The media/platform/website reporting the disinformation (SOURCE):

informiran.net

The link to the disinformation (URL):

https://www.informiran.net/%D0%BB%D0%B5%D0%BD%D0%B0-
%D0%B1%D0%BE%D1%80%D0%B8%D1%81%D0%BB%D0%B0%D0%B2%D0%BE%D0%B2%D0%B0-
%D0%BD%D0%B0-%D1%88%D0%B8%D0%BF%D0%BA%D0%B0-%D0%B8%D0%BC%D0%B0-
%D0%BA%D0%BE%D1%81%D1%82%D0%BD%D0%B8%D1%86/

Language of the claim (LANGUAGE):

Bulgarian

The textual statement of the disinformation and analysis in English and in the source language (TEXT),
(TEXT:SOURCE LANGUAGE):

Лена Бориславова: На Шипка има костница, да ги отделим от другите, защото не може руски кости да лежат в българската земя

Translation: **Lena Borislavova: There is an ossuary in Shipka, let's separate them from the others, because Russian bones cannot lie on Bulgarian soil**

The fact-checking analysis (WHY):

Bulgarian politician Lena Borislavova, a frequent target of hate campaigns, is quoted as saying that the bones in the Shipka memorial of Russian soldiers who lost their lives during the Russian-Turkish war of 1877-78 should be separated from the bones of Bulgarians. In fact only the title links the commentary to Borislavova. The quote is in fact part of a commentary by Desislava Stancheva (possibly a pen name). When the publication is shared on social media, the fake quote makes appear that Borislavova makes the scandalous comment.

The fact-checking analysis in the source language (WHY:SOURCE LANGUAGE):

Българският политик Лена Бориславова, чест обект на кампании на омраза, е цитирана да казва, че костите в паметника на руските войници на Шипка, загинали по време на Руско-турската война от 1877-78 г., трябва да бъдат отделени от костите на българите. Всъщност само заглавието свързва коментара с Бориславова. Всъщност цитатът е част от коментар на Десислава Станчева (вероятно псевдоним). Когато публикацията е споделена в социалните мрежи, фалшивият цитат създава впечатление, че Бориславова прави скандалния коментар.

The audios, videos, images related to the disinformation (MULTIMEDIA):

URL of Png/jpeg file extracted from the news

Rating scale used to assess its truthfulness (RATING SCALE):

Fake

Country of the disinformation (COUNTRY):

Bulgaria

FIG.2 EXAMPLE OF INFORMATION EXTRACTION FROM A DISINFORMATION CASE

2.2 DATABASE MANAGEMENT SYSTEM
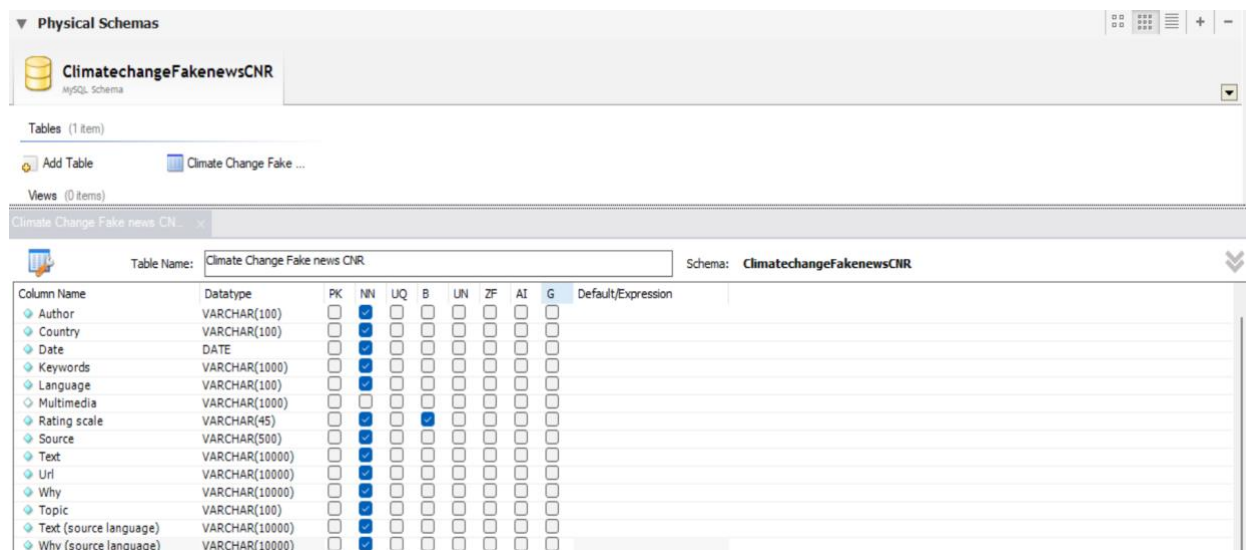
To identify the reference database management system (DBMS) used for the implementation of the two datasets, the study provided by Taipalus (2023) has been analysed, which provides a comparison of the performances of different DBMSs. The *MySQL (https://www.mysql.com/)* DBMS has been selected mainly due to its overall performance, scalability, and reliability. As an open-source DBMS, *MySQL* offers a cost-effective solution without any significant compromising on features or support. Its compatibility with a wide range of platforms and programming

languages ensures a good level of integration with the existing technology stack. Moreover, it allows the sharing of data in accessible and common file formats.

The active users and developers' communities around *MySQL*, along with extensive documentation, provide ample resources for troubleshooting and optimization. Furthermore, *MySQL*'s high availability and data security features, such as replication and automatic backup, ensure data integrity and minimal downtime, making it a good choice for mission-critical applications. These factors, combined with *MySQL*'s proven track record in managing large-scale and high-transactional databases, made it the preferred choice for this project.

For each of the two case studies of the project (war in Ukraine and climate changes), a specific dataset model, based on the template described in Section 2.1, was designed. Each dataset was built around a single table, where each claim has been associated with all the fields already present on the approved template (see Fig. 2).

Fig. 3 and Fig. 4 provide some visual examples of the dataset model and table structure.

FIG. 3 DATASET MODEL STRUCTURE

idtable1	Author	Country	Date	Keywords	Language	Multimedia	Rating scale	Source	Text	Url	Why	Topic	Text (source language)	Why (source language)
1	N/A	Poland	12/03/2024	vaccinations, cli...	Polish	https://drive.go...	Fake	Demagog	A video in which ...	https://demago...	There is no evid...	Climate change	Na Facebooku p...	Nie ma dowodó...
2	N/A	France	01/08/2022	climate protest, ...	French	https://drive.go...	Fake	AFP Factuel	A video clip of m...	https://factchec...	This clip shows a...	Climate change	Un clip vidéo d'a...	Ce clip montre u...
3	N/A	Lithuania	23/04/2022	video, climate, ...	Lithuanian	https://drive.go...	Fake	Delfi	The video, shar...	https://www.del...	Video of anti-cli...	Climate change	„YouTube" ir kitu...	Vaizdo įrašas api...
4	N/A	Germany	22/03/2022	body bags, clima...	German	https://correctiv...	Fake	Correctiv	This video doesn...	https://correctiv...	The video recor...	Climate change	Dieses Video zei...	Die Videoaufnah...
6	Xavier Trias	Spain	31/05/2024	cars, pollution	English	N/A	Fake	eufactcheck-eu	Cars are related ...	https://eufactch...	As we have see...	Climate change	Cars are related...	As we have see...
7	Neva Zganec, Ni...	Croatia	16/04/2024	Croatia, climate ...	Croatian	N/A	Fake	eufactcheck-eu	Migrants are ma...	https://eufactch...	There is no conn...	Climate change	Migranti masovn...	Ne postoji veza i...
8	Donald Trump	USA	25/12/2023	global warming, ...	English	https://drive.go...	Fake	eufactcheck-eu	The concept of ...	https://eufactch...	We haven't foun...	Climate change	The concept of ...	We haven't foun...
9	Stephan Brander	Germany	27/05/2023	Germany, wind ...	German	N/A	Fake	eufactcheck-eu	Germany will ne...	https://eufactch...	Stephan Brande...	Climate change	Deutschland wir...	Stephan Brande...
10	Vivek Ramaswamy	USA	15/092023	Ramaswamy, Cli...	English	N/A	Fake	FactCheck.org	A former biotech...	https://www-fac...	Republican presi...	Climate change	A former biotech...	Republican presi...
11	Steve Milloy	USA	26/01/2023	Global Temperat...	English	https://drive.go...	Fake	Factcheck.org	Steve Milloy, wh...	https://www-fac...	That's wrong. N...	Climate Change	Steve Milloy, wh...	That's wrong. N...
12	Donald Trump	USA	05/09/2020	Climate, Wildfire...	English	N/A	Fake	Factcheck.org	In a briefing on t...	https://www-fac...	The Earth is war...	Climate change	In a briefing on t...	The Earth is war...
13	Donald Trump	USA	12/12/2019	Global Warming,...	English	N/A	Fake	Factcheck.org	At a campaign r...	https://www-fac...	According to the...	Climate change	At a campaign r...	According to the...
14	Kamala Harris, E...	USA, Fra...	22/08/2019	Amazon, World, ...	English	N/A	Fake	Factcheck.org	The Amazon pro...	https://www-fac...	No. Scientists es...	Climate change	The Amazon pro...	No. Scientists es...
15	Jørgen Peder St...	Germany	29/08/2023	Global Warming,...	German	N/A	Fake	Correctiv-org	Jørgen Peder St...	https://correctiv...	Steffensen's sta...	Climate change	Jørgen Peder St...	Steffensens Aus...
16	N/A	Germany	09/02/2024	Floods, Water le...	German	https://drive.go...	Fake	Correctiv-org	A photo with a h...	https://correctiv...	There have also ...	Climate change	Ein Foto mit eine...	Fehlender Konte...
17	N/A	Germany	25/01/2024	Wind turbines, T...	German	https://drive.go...	Fake	Correctiv-org	On Facebook, u...	https://correctiv...	A nuclear power...	Climate change	Auf Facebook m...	Ein Atomkraftwe...
18	Heute Journal	Germany	21/12/2023	Global Warming	German	N/A	Fake	Correctiv-org	In a news report...	https://correctiv...	The map was no...	Climate change	Das ZDF habe in...	Manipuliert. Die ...
19	N/A	Germany	21/12/2023	E-cars, prices fo...	German	https://drive.go...	Fake	Correctiv-org	The electricity c...	https://correctiv...	The price inform...	Climate change	Der Strom-Lades...	Die Preisangabe ...
20	N/A	Germany	30/10/2023	Ukraine, War CO2	German	N/A	Fake	Correctiv-org	The University o...	https://correctiv...	Heidelberg Univ...	Climate change	Die Universität H...	Die Universität H...
21	Robert Farle	Germany	23/04/2023	Germany, CO2 e...	German	N/A	Fake	Correctiv-org	78 percent of ou...	https://correctiv...	It is true that th...	Climate change	78 Prozent unse...	Es stimmt, dass ...
22	N/A	Germany	26/06/2023	Wind farms, dry...	German	https://drive.go...	Fake	Correctiv-org	According to a C...	https://correctiv...	The study show...	Climate change	Laut einer chine...	Die Studie zeigt,...
23	N/A	Germany	22/06/2023	Heat protection ...	German	N/A	Fake	Correctiv-org	The next lockdo...	https://correctiv...	Health Minister K...	Climate change	Die nächsten Lo...	Gesundheitsmini...
24	N/A	Germany	07/06/2023	Heat wave	German	N/A	Fake	Correctiv-org	According to a s...	https://correctiv...	Kachelmann's tw...	Climate change	Laut einem Shar...	Der Tweet von K...
25	N/A	Germany	07/06/2023	Cold, Antarctica,...	German	https://drive.go...	Fake	Correctiv-org	A new cold recor...	https://correctiv...	It is true that th...	Climate change	Ein neuer Kälter...	Richtig ist, dass ...
26	N/A	Germany	15/09/2022	Wind turbines, A...	German	https://drive.go...	Fake	Correctiv-org	Photos show tha...	https://correctiv...	The forest area ...	Climate change	Fotos belegten, ...	Die Waldfläche b...
27	N/A	Germany	09/09/2022	Trees, Wind tur...	German	N/A	Fake	Correctiv-org	To save the cou...	https://correctiv...	The trees were ...	Climate change	Um das Land vor...	Die Bäume wurd...
28	Harald Laatsch	Germany	08/09/2022	Alpine glaciers, ...	German	N/A	Fake	Correctiv-org	A simulation of t...	https://correctiv...	The retreat of t...	Climate change	Eine Simulation d...	Der Rückgang d...

FIG. 4 SAMPLE OF DATASET TABLE

2.3 CLOUD STORAGE AND DATA SHARING

All multimedia files, extracted from the reference fact-checking websites and the existing databases owned by EURACTIV Bulgaria and Internews Ukraine, have been stored online instead of being directly inserted into the database to have more resources (e.g. computation and storage capabilities) for executing queries efficiently and making data processing and analytics more efficient. Free versions of two online file storage services were used that allow maintaining high-performance levels of the datasets over time and making smoother the tasks of subsequent WP that will use the datasets as input.

The cloud storage services used were *Mediafire* (https://www.mediafire.com) and *Google Drive* (https://www.google.com/intl/it_it/drive/).

These services were also used to store backup files and to improve teamwork efficiency and online collaboration by taking advantage of the following benefits and features:
- **Scalability**: online storage provides a flexible and scalable infrastructure that allows to expand storage capacity on demand. This is particularly advantageous for DBMS, which often needs to handle growing amounts of data.
- **Accessibility**: users can access the database from anywhere with an Internet connection, facilitating remote work and global collaboration. This real-time collaboration led to more efficient workflows and better productivity.

- **Data Security**: online storage providers often implement strong security measures, including encryption, firewalls, and intrusion detection systems. These measures can be regularly updated compared to what many businesses can afford to implement on their own on-premises systems.
- **Recovery Solutions**: online storage provides reliable disaster recovery options. Data stored in the cloud can be quickly restored in the event of hardware failures, natural disasters, or cyberattacks. Cloud providers often offer automated backup solutions that ensure data is regularly backed up and can be recovered with minimal downtime.

The first version of each dataset is delivered in the following file formats:
- A CSV file with semicolon separator.
- The model file (.mwb), containing the physical schema of the dataset and the populated table.

These file formats allow to easily edit, process, and manage data, with also the possibility to connect the database models to internal or external servers. While the model structure was already shown in section 2.2 (Fig.3), Fig.5 shows a sample portion of an exported CSV file.

FIG 5. SAMPLE OF EXPORTED CSV FILE

3   DISINFORMATION CASES COLLECTED IN THE FIRST RELEASE

In the first release of the datasets, 1000 disinformation cases have been collected: 577 cases related to the war in Ukraine and 423 cases related to climate change.

In the following sub-sections, a brief description of the disinformation cases on the war in Ukraine and climate changes extracted by CNR, Euractiv Bulgaria, and Internews Ukraine are illustrated.

3.1   DISINFORMATION CASES ON THE WAR IN UKRAINE

3.1.1   DISINFORMATION CASES EXTRACTED BY CNR

The spreading and proliferation of fake news about the war in Ukraine within the European Union has become a critical issue, which deeply impacts both public perception and policymaking. These false narratives are often crafted through disinformation campaigns, utilizing social media platforms and other digital channels to spread misleading or entirely falsified information. The main goal of this misinformation is to manipulate public opinion and undermine support for Ukraine. To understand the scope and impact of this phenomenon, CNR has carefully collected and analysed 300 claims related to the conflict. This comprehensive collection aims to highlight the methods and motivations behind these falsehoods, providing valuable insights into the mechanisms of modern propaganda and its effects on European societies.

To do this, the analysed claims were extracted from certified and recognized fact-checking websites. In this regard, the websites have been selected from the repository of the European Digital Media Observatory (EDMO). In particular, the selected fact-checking websites were:

- correctiv
- demagog
- demagog.cz
- logically facts
- facta.news
- tjekdet.dk
- efe verifica
- afp factuel
- factcheck vlaanderen
- dpa-factchecking.com

The main collected topics were concerned with four disinformation misleading narratives about the war in Ukraine:

- from the Third World War to a biological disaster: a widely recognized instance of disinformation circulating throughout Europe is designed to provoke panic and spread

fear among the public. This is primarily achieved through two frightening assertions: that additional nations are prepared to enter the conflict, and that the war has the potential to trigger a catastrophic biological or nuclear disaster.

- Zelensky and Ukrainian army: disinformation has targeted the Ukrainian army and the country's president, Volodymyr Zelensky. Misleading content has depicted him in contrasting ways. From a pro-Russian viewpoint, he is portrayed as a nazi and a coward, while from a pro-Ukrainian perspective, he is seen as a national hero.
- questioning the authenticity of the conflict: another popular disinformation narrative claims that the Ukrainian conflict is not real.
- disinformation about refugees: false claims about millions of refugees attempting to flee Ukraine and seek refuge in neighbouring European countries (primarily Poland, but also Hungary, Slovakia, and Romania) were identified according to the United Nations, by March 8, 2022, false claims have been made that over 2.3 million people had left Ukraine, with 1.4 million entering Poland.

3.1.2 DISINFORMATION CASES EXTRACTED BY EURACTIV BULGARIA

Disinformation cases regarding the Russian aggression against Ukraine are commonplace in Bulgarian media, to a much greater extent compared to other EU countries. The reasons for this are severalfold. In Bulgaria, many media outlets are opaque regarding their ownership and are not accountable as to their content. The Bulgarian media regulator has a '*laisser-faire*' attitude to media promoting Russian messages, considering that "all opinions must be heard". Among the Bulgarian political parties, several are sympathetic to the Kremlin (the party "Vazrazhdane" is openly pro-Russian, the Bulgarian Socialist Party is sympathetic to the Russian official positions, and several smaller parties are repeating the messages propagated from Moscow). As a result, the Bulgarian public opinion is divided regarding the war in Ukraine. Bulgaria is the EU country with the highest percentage of people who have a positive opinion of Vladimir Putin (37%)[1].

Efforts in Bulgaria to debunk the fake news are made by fact-checkers supported by Western financing, but their activity has not discouraged the many media outlets churning disinformation to continue their activity.

Disinformation takes place in various forms. A very common form is to translate propaganda articles from the Russian state-supported press without indication of source. Another is to mix news with interpretations and to assign a piece of information a completely misleading title.

Many commentaries intended to disinform the audience are not signed, or the signature is a pen name, which doesn't seem to matter to the readers. In most cases, the shocking title is intended to be shared on Facebook, the dominant social media in the country. Research has shown that

---

[1] https://www.euronews.com/my-europe/2024/04/05/which-eu-countries-like-and-dislike-zelenskyy-the-most-euronews-poll-shows-big-swings-in-o#:~:text=The%20%22negative%22%20rate%20falls%20under,highest%20mark%20across%20the%20board.

most Bulgarians get their information from Facebook and that they rarely click on the article, considering that the title is enough for them. For the scope of the Ai4Debunk project, Euractiv Bulgaria has carefully collected and analysed 127 claims related to the conflict.

Among the favourite topics/narratives of disinformation are:

- Ukraine wants to push NATO toward a full-scale war with Russia.
- NATO wants to have bases on the soil of Bulgaria and can bring all kinds of weapons in total secrecy, including nuclear weapons.
- NATO will push for conscription.
- NATO will send troops to Ukraine, including Bulgarians.
- Ukraine is corrupted: many weapons and equipment sent to Ukraine get stolen.
- Ukrainian President Volodymyr Zelensky is a clown, a drug addict. He is extremely rich and gets even richer as the war goes by.
- The political parties in Bulgaria that support Ukraine are either stupid or corrupt, or both. Their political leaders have secret agendas and are ready to sacrifice young Bulgarians on the front in Ukraine just to get a tap on the back from Washington.
- The US is pressuring Bulgaria to give Ukraine the two Soviet-made nuclear reactors intended for a nuclear plant that will never be built free of charge (they cost billions).
- Bulgaria is losing billions because of the duty-free imports of cheap cereals from Ukraine authorised by the EU.
- The West (the CIA) has plans to eliminate the leaders who are critical of the Western stance on Ukraine; the assassination attempt against Robert Fico being highlighted as the proof of this thesis.

The list is not exhaustive and new topics appear more and more often.

3.1.3   DISINFORMATION CASES EXTRACTED BY INTERNEWS UKRAINE

Russian media have been disseminating a series of disinformation campaigns in online media and Telegram channels related to the War against Ukraine, targeting various aspects to manipulate public perception and further geopolitical agendas. The 150 cases present various instances of disinformation spread by Russian media and propaganda during the war in Ukraine since the full-scale Russian invasion. It provides specific examples of false claims made about Poland annexing Western Ukraine, American Biolabs operating in Ukraine, forced mobilization of women, and Ukrainian military involvement in illegal weapon trade. The text also highlights the destruction and war crimes committed by Russian forces in cities like Mariupol and Bucha, as well as the exploitation and deportation of Ukrainian children by Russia. The disinformation is debunked with evidence and statements from reliable sources.

The main topics covered in the disinformation cases related to the War in Ukraine can be categorized as follows:

1. Territorial Integrity and Sovereignty

- Poland Occupying Western Ukraine

The false claim that Poland is planning to annex Western Ukraine with the approval of the EU and USA aims to suggest that Ukraine's neighbours have intentions of reclaiming its territory.

2. Russia as peacekeepers

On January 30, 2024, ZOV Mariupol and REN TV claimed that people in Germany were impressed by the speed of Russia's rebuilding efforts in Mariupol. This report manipulated the narrative, as Russia's siege had caused massive destruction, and post-occupation efforts were more about demolishing damaged buildings and erasing Ukrainian cultural identity. This was rated as manipulation.

3. Military and security threats
   - US Biological Weapons in Ukraine

The narrative that the USA has been developing biological weapons in Ukraine, portraying the country as a significant threat to Russia and global security.

- NATO Involvement in Attacks on Russia and NATO's mercenaries

Misrepresentation of the NATO Secretary General's statements to falsely claim that NATO admitted involvement in attacks on Russia, suggesting direct military engagement by NATO countries.

On March 7, 2024, Dmitry Kaverin from gazeta.ru falsely claimed that Ukraine's Main Intelligence Directorate admitted to the involvement of NATO mercenaries in the war. This was misleading, as the Directorate had referred to international volunteers who are legally participating in Ukraine's military operations, abiding by both Ukrainian and international laws. This claim was detected and rated as fake by Ukraine`s Center for Strategic Communications and Informational Security. Later, Ukrainian national media (e.g. 1+1) also reported this fake.

- Military support that Ukraine received from partners

On July 5, 2022, Boyko Nikolov from bulgarianmilitary.com asserted that Ukraine had sold two French-donated 155mm Caesar howitzers to Russia. This claim, based on unreliable sources, was debunked by the French military. It was intended to paint Ukraine as corrupt and untrustworthy, thereby discouraging Western support. This too was rated as fake.

4. Human rights violations and internal governance
   - Concentration Camps for Avoiding Conscription

The baseless allegation that the Ukrainian government has established concentration camps for men avoiding conscription intends to depict Ukraine as an oppressive country violating human rights.

- EU deporting Ukrainian men for conscription

The claim that the EU is planning to deport Ukrainian men for conscription purposes, is aiming to show EU actions as coercive and harmful towards Ukrainian refugees.

- Mobilization in Ukraine for pregnant women

Another instance occurred on February 10, 2024, when Tatiana Bielova from smotrim.ru reported that Ukraine had started the forced mobilization of women, including pregnant women. This was incorrect; the mobilization of women in Ukraine is voluntary and limited to those with specific medical or pharmaceutical qualifications. The false report aimed to depict Ukraine's mobilization efforts as desperate, intending to demoralize the population. This was also rated as a fake.

5. Political legitimacy
   - Zelensky's legitimacy and Western abandonment

The narrative that President Zelensky would become illegitimate if elections were not held due to martial law, suggesting a potential loss of Western support and questioning his continued authority.

6. War Crimes committed by Ukraine

On June 6, 2023, a Telegram channel called "War with fakes" and ZOV Kherson falsely claimed that the Ukrainian Armed Forces destroyed the Kakhovka HPP. Evidence suggested that the explosion was initiated from the Russian-controlled side, not by Ukrainian forces.

Aleksandr Khristenko from Vesti.Ru, on April 4, 2022, propagated the false narrative that the Bucha massacre was staged by Ukrainians. Substantial evidence and witness accounts have documented Russian war crimes in Bucha, including executions and unlawful detentions.

Further disinformation was propagated on April 5, 2022, by Nikita Makarenkov and Pavel Khanarin from KP.Ru Donetsk. They alleged that Ukrainian forces ordered the annihilation of witnesses to nationalist crimes in Mariupol, with plans to blame Russian and DPR forces for civilian deaths. In reality, Russian forces had committed numerous attacks on civilian infrastructure during the siege of Mariupol, not the Ukrainian military. This claim was rated as fake.

These topics reflect the overarching goals of the disinformation campaigns: undermining Ukraine's sovereignty, portraying it as a threat or a failing state, questioning the legitimacy of its government, and distorting the nature of international support for Ukraine.

3.2   DISINFORMATION CASES ON CLIMATE CHANGE

3.2.1   DISINFORMATION CASES EXTRACTED BY CNR

The dissemination of false claims about climate change alongside the European Union sets a significant challenge to both public awareness and policy efforts aimed at mitigating environmental impact. These deceptive narratives are often shared through digital media, by using social networks and websites to spread misinformation discrediting the scientific consensus

on climate change. The motivations behind these efforts range from political agendas to economic interests, with the attempt to delay or obstruct climate action. To address this issue, CNR has collected 400 claims related to climate change misinformation. This extensive collection not only highlights the pervasive nature of false information, but also aims to uncover the main strategies adopted in these disinformation campaigns to strengthen public understanding and resilience against these misleading narratives.

For this case study, the analysed claims were extracted from certified and recognized fact-checking websites. The websites have been mainly selected from the repository of European Digital Media Observatory (EDMO) and eufactcheck.eu.
In particular, the selected fact-checking websites were:
- eufactcheck.eu
- factcheck.vlaanderen
- Faktisk.nomimikama.org
- ellinikahoaxes.gr
- correctiv
- demagog
- logically facts
- facta.news
- afp factuel
- factcheck vlaanderen

Moreover, considering the specific topic of this case study, further specific feedback coming directly from the scientific community was collected. To achieve this goal, several claims have been extracted from the following fact-checking websites:
- skepticalscience
- science.feedback.org

The main collected topics were concerned with the following misleading narratives:
- solar activity and climate changes: according to a high number of false claims, the long and short-term variation in solar activity would play a high role in climate changes.
- CO_2 role in climate changes: according to several claims shared through newspapers and social media platforms, CO_2 would not have a provable effect on the climate, the greenhouse effect is negligible, and computer models used to measure climate change would be skewed.
- global warming and ice melting: numerous false claims about rising sea levels and climate warming were shared.

- agriculture, climate, and green deal: there have been various false claims and misunderstandings circulating about farmers and the EU's green deal, particularly relating to its impact on agriculture and the livelihood of farmers. A prevalent falsehood is that the Green Deal will decimate farming in Europe.
- Greta Thunberg's activity against climate change: some fabricated false news about Greta Thunberg's declarations on climate change was spread on the web. In this regard, an example comprehends fake news articles that falsely report that Greta Thunberg declared climate change to be a fabricated issue, created by politicians to control the public, admitting also that her climate activism is merely a publicity stunt for fame and financial gain.
- Vulcan activity and climate changes: some false claims stated that the eruption of a single volcano can offset centuries of anthropogenic impact on climate change.
- chemical trails, geoengineering, climate, and health: numerous false news state that governments are banning chemtrails, confirming secret government programs using chemtrails for various purposes that would harm the public and the environment.
- electric vehicles, ecology, climate, and pollution: according to some false information circulating on the web and social media, electric vehicles would be worse for the climate than gasoline cars because of power plant emissions. Also, electric vehicles would have a higher carbon footprint than gasoline cars, even for the electricity used for charging.

3.2.2   DISINFORMATION CASES EXTRACTED BY EURACTIV BULGARIA

Climate change in Bulgaria is a big topic in the anti-EU narrative that has become plain to see on the campaign for the EU elections. In Bulgaria, there are 20 political parties and 11 coalitions running for elections, and most of these political players are entitled to state funding through media packages. Traditionally, funds for media advertising are the main expense in election campaigns, and political entities are free to choose where to advertise.

Most media take the money gladly, and state broadcasters, radio, and TV, are legally obliged to give equal advertising space to all political players. It matters little if out of 20 parties, only five can realistically make it to Parliament – all are equally benefitting from the media to voice their messages.

The average citizen is exposed to a significant amount of anti-EU rhetoric, making it remarkable that some voters continue to support political forces that do not seek to dismantle the EU.

An analysis of the electoral messages reveals that the European Green Deal is the primary target of criticism for the majority of these 20 parties and 11 coalitions.

And those openly criticizing the flagship policy combining environmental and climate action of the outgoing European Commission are not only fringe parties. They include the Bulgarian Socialist Party (BSP), a PES and S&D member who says the European Green Deal, adopted in 2020, should be renegotiated.

In this context, Euractiv Bulgaria has collected 23 claims analysing the press churning messages hostile to the EU's climate policies. Among them are:

- The EU wants Bulgaria to lose its base load capacities (nuclear and coal), and Brussels wants Bulgaria, a traditional exporter of electricity, to become an importer of electricity.
- Wind turbines or solar energy cannot replace the traditional baseload capacities, they are useless when there is no wind or sun.
- The European Commission is led by "green Talibans" who need to be replaced.
- The Green Deal must be renegotiated or abolished.
- Climate change is a hoax, or at least it is part of the "politically correct" policymaking embodied by the outgoing European Commission that is coming to an end.
- Bulgaria should not reduce its CO2 emissions because it's a very green country, Instead, Bulgaria should receive direct payments for preserving its forests.
- With Donald Trump back in the White House, the climate change policies will be reconsidered anyway.

4 DISINFORMATION CASES COLLECTED IN THE SECOND RELEASE

In the second release of the datasets, 550 disinformation cases were collected: 237 cases were related to the war in Ukraine and 313 cases were related to climate change.

In the following sub-sections, a brief description of the disinformation cases on the war in Ukraine and climate changes extracted by CNR, Euractiv Bulgaria, and Internews Ukraine are illustrated.

4.1 DISINFORMATION CASES ON THE WAR IN UKRAINE

4.1.1 DISINFORMATION CASES EXTRACTED BY CNR

For the second release of the dataset, CNR has continued to follow the latest insight and early warning on disinformation narratives about the war in Ukraine. To continue evaluating and understand the extent of this misinformation, CNR has collected and analysed further 50 claims related to this topic. The approach used in this selection was once again designed to emphasize

the techniques and intentions behind these falsehoods, offering relevant insights into the workings of contemporary propaganda and its impact on European and global scenarios.

Even for the second release of the dataset, the analysed claims were extracted from certified and recognized fact-checking websites. Once again, the websites have been selected from the repository of the European Digital Media Observatory (EDMO). In particular, the selected fact-checking websites were:

- Correctiv
- knack
- demagog
- demagog.cz
- logically facts
- facta.news
- tjekdet.dk
- efe verifica
- afp factuel
- factcheck vlaanderen
- dpa-factchecking.com

The main collected topics were concerned with five disinformation misleading narratives about the war in Ukraine:

- Counter-offensive related disinformation: The Ukrainian counter-offensive is a failure.
- Discrediting the Ukrainian government: Ukrainian authorities are corrupted and embezzling international aid funds; Damaging Zelensky's public figure.
- Compromising Ukrainians' image: Ukrainians are lying about the war; Ukrainians are Nazis and murderers; Ukrainian refugees are a danger to host societies.
- Legitimizing Russia and the invasion: Russia has many powerful allies; Russia is fighting a war against NATO/The West.
- Attacking NATO and Ukraine foreign supporters: Economic support to Ukraine is backfiring on EU populations; Ukraine used to revamp conspiracy theories against the West

4.1.2 DISINFORMATION CASES EXTRACTED BY INTERNEWS UKRAINE

Russian media have been disseminating a series of disinformation narratives related to the War against Ukraine in online media and social media accounts (in X, TikTok, Facebook, and Instagram). The gathered 150 cases present diverse manifestations of Russian disinformation and propaganda spread during the war against Ukraine since the beginning of the full-scale invasion of Ukraine in 2022. All the 150 cases of Russian disinformation and propaganda are represented

in diverse formats — textual statements, photo and video fakes. They also provide examples of Russian disinformation and propaganda narratives targeting internal audiences in Ukraine (e.g. narratives related to mobilization in Ukraine, social and political situation in Ukraine, narratives undermining trust in the Ukrainian authorities, narratives targeting Ukrainian allies in the EU countries, and in the USA, gender disinformation) as well as external audiences, in particular in the EU countries (e.g. narratives discrediting Ukrainian refugees in the EU, narratives discrediting Ukrainian authorities, narratives about sending of soldiers from the EU countries to fight in Ukraine, narratives about black market of Western weapons, human organs and children`s trafficking from Ukraine). To debunk these cases, evidence and statements from reliable sources and fact-checking resources were used as well as analysis of visual content (reverse search of images, using special tools to detect traces of photo editing).

Generally, key topics covered in disinformation cases related to the War in Ukraine can be categorized in the following way:

1. Mobilization in Ukraine

Russian Telegram channels disseminated narratives related to mobilization in Ukraine. Mostly, these messages were disseminated in Telegram channels (e.g. Sheikh Tamir, GASPARYAN, Petya Perviy, Cherniy Kvartal). They can be categorized in the following categories:

- **Mobilization of people with disabilities or mental diseases —** on 10 May pro-Russian channel operating in occupied territories of Ukraine "Ryadovoy na Peredovoy" published a staged video with a man with Down syndrome taking military service on the frontline and facing humiliation from other soldiers. Similar videos were also published in other pro-Russian Telegram channels, like Stepnoy Veter, involving the same plot and the same story of the man with Down Syndrome taking military service.

Similarly, on 5 June on the Russian Telegram channel, a fake photo appeared from Mykolaiv Mental Hospital of the announcement to relatives of patients to come to the hospital during the work of the military medical commission.

- **Mobilization of students or involving them in military works —** on 26 July TikTok account published a deepfake video with the Kyiv mayor Vitaliy Klychko stating that students must be mobilized because of their laziness.

Similar messages were published on 19 June on the Telegram channel GASPARYAN — there was published a fake photo from Kyiv University where it was stated that students must go to dig trenches instead of practice.

- **Refusal of state or commercial companies to provide their services because of work of Territorial Centres of Recruitment and Social Support —** on 12 June Russian Telegram Channel Sheikh Tamir published a message that workers of ambulance in Odesa refuse to come to calls to districts where workers of Territorial Centres of Recruitment and Social Support operate using the fake photo of the announcement from the hospital.

- **Mobilization of women —** on 4 July Russian Telegram Channel Sheikh Tamir published a video with claims that workers of Territorial Centres of Recruitment and Social Support began to mobilize women in Ukraine. In the video, which was originally taken on 30 June during the conflict between a woman and border guards, it is shown that a woman has a conflict with men in military uniform. It was used as evidence of the "mobilization of women in Ukraine'.

2. Ukraine as a "terrorist state"

On 8 July 2024, Russia launched an attack on Kyiv hitting the children`s hospital Okhmadyt. This event was used by Russian and pro-Russian bloggers and journalists to accuse Ukraine of committing an attack on civilian infrastructure. For instance, on 8 July Russian media Tsargrad published news with the claim that it was the fault of the Ukrainian air defence system, and the hospital was hit by the US missile AIM-120 AMRAAM from NASAMS.

Similarly, on 8 July Russian Telegram channel Perviy Kharkovskiy disseminated messages about the Ukrainian air defence system hitting the residential building and using the children`s hospital as a "military object" since the beginning of the full-scale invasion.

3. Military and security threats

- **Involvement of foreign mercenaries in combat actions in Ukraine**

On 4 May Asia Times disseminated the news about the involvement of the French legionary in combat actions in the Donetsk region to support the AFU. These statements were based on claims from Russian media Sputnik and Russia`s Pravda.

- **NATO Mercenaries in the AFU**

On 2 May, the Telegram channel of the Russian propagandist Dmitry Vasilets disseminated the video with fake statements of the AFU soldiers about NATO commanders executing and humiliating AFU soldiers. The video, that was used as evidence, appeared to be a part of the interview with the former Russian soldiers of the PMC "Wagner" taken on 20 March. In this interview he talked about the realities in the PMC "Wagner" however, his words were presented as the interview of the AFU soldier talking about "NATO commanders".

- **Narratives undermining the Western humanitarian aid to Ukraine**

On 22 July Telegram Channel Sheikh Tamir disseminated fake news about Nestle sending water infected by fecal bacteria to Ukraine and Palestine as humanitarian aid. The fake screenshot from the company's X-account was used as evidence, where it was stated that the company stops to sell its products and sent water to Ukraine and Palestine.

Similarly, on 5 November 2023, Russian resource Ukraina.Ru published fake news about expired medicine that was sent to Kharkiv and Dnipro as humanitarian help and led to the deaths of patients.

- **"Black market of weapons, human organs and children`s trafficking"**

On 22 July Russian Foundation to Battle Injustice published research accusing Ukraine of trading human organs, Western weapons, and children with the Italian mafia. Similar articles appeared in other pro-Russian media targeting foreign audiences. The research is based on words of pro-Russian "experts", anonymous sources and visuals taken from a website that also has links to Russia.

4. Ukrainian refugees in the EU

Russian and pro-Russian channels disseminated fake news about Ukrainian refugees being involved in the testing of cosmetics by French companies on Ukrainian female refugees (Telegram channel "Ukropskiy Fresh" on 8 June) and Ukrainian refugees becoming the biggest kidney donor in the EU (Telegram channel "KIBERFRONT ZA Rossiyu", 16 August). All these fake statements contained videos that were presented as videos from the reliable EU media, but after the analysis it appeared to be edited.

4.1.3 DISINFORMATION CASES EXTRACTED BY EURACTIV BULGARIA

Euractiv Bulgaria has continued to follow the latest insight and early warning on disinformation narratives about the war in Ukraine by collecting and analysing further 15 claims related to this topic.

The war in Ukraine is portrayed through a pro-Russian lens, with headlines like *"Six battalions of the Armed Forces of Ukraine refused to carry out a combat mission"* and *"Climate is not what it seems: The US is planning a war near Russia's borders"* undermining Ukraine's efforts and casting the West as the aggressor. Claims that Bulgaria is being forced into supporting Ukraine, like *"We give it [the Belene nuclear reactor, a Soviet-built pressure vessel Bulgaria is unable to use] to Ukraine"*, further feed into fears of national exploitation and foreign intervention and connects the narratives of energy and national security. This narrative builds on historical and cultural ties with Russia, presenting Ukraine as a corrupt or failed state and amplifying distrust in Western actions, whether through NATO or EU energy policies.

These disinformation campaigns aim to erode public trust in international alliances, particularly Bulgaria's position within the EU and NATO. By creating a sense of national victimhood, where Bulgaria is seen as being manipulated or harmed by external forces, these narratives strengthen isolationist sentiments and weaken support for broader cooperative efforts on both environmental and geopolitical issues. Ultimately, this widespread disinformation fosters polarization, harms EU's integrity, delays necessary climate action and energy transition, and undermines solidarity with Ukraine in its war with Russia.

4.2   DISINFORMATION CASES ON CLIMATE CHANGE

4.2.1   DISINFORMATION CASES EXTRACTED BY CNR

A high number of false claims about climate change are continuing to circulate through digital platforms, in particular through social networks and websites to undermine the established scientific consensus on climate change. To address this issue, CNR has collected further 200 claims related to climate change misinformation for the second release of the database. This extensive data collection not only highlights the pervasive nature of false information, but also aims to uncover the main strategies adopted in these disinformation campaigns to strengthen public understanding and resilience against these misleading narratives.

Once again, the analysed claims were extracted from certified and recognized fact-checking websites. The websites have been mainly selected from the repository of the European Digital Media Observatory (EDMO) and eufactcheck.eu.

In particular, the selected fact checking websites were:

- eufactcheck.eu
- Faktisk.nomimikama.org
- correctiv
- facta.news
- afp factuel
- factcheck vlaanderen

Even for the second release of the dataset, considering the specific topic of this case study, further specific feedback coming directly from the scientific community and global fact-checking platforms was collected. To achieve this goal, several claims have been extracted from the following fact-checking websites:

- skepticalscience
- the daily sceptic
- The epoch times
- notrickszone.com
- Politifact.com
- Kritichno.bg
- travelask.ru
- science.feedback.org

The main collected topics were concerned with the following misleading narratives:

- Wind turbines: Numerous false claims about alleged damages to the environment caused by wind turbines have been collected.
- Conspiracy Theories: Allegations that climate change is a hoax orchestrated by governments, scientists, or special interest groups to control the public or advance hidden agendas.
- Misrepresentation of Data: Misleading statistics, cherry-picking data, or outdated information are used to support claims that undermine the reality or severity of climate change.
- Scepticism of Scientific Consensus: Arguments that there is no agreement among scientists about climate change, suggesting that the science is unsettled or controversial.
- Denial of Climate Change: Claims that climate change is not happening or that the Earth's climate has always changed naturally without human influence.
- Discrediting Renewable Energy: Spreading false information about the inefficacy, costliness, or impracticality of renewable energy sources, often to support fossil fuel industries.
- Misleading Narratives About Policy Impacts: Exaggerating or falsifying the negative impacts of climate policies, such as job losses or energy shortages, to create resistance against climate legislation.

4.2.2 DISINFORMATION CASES EXTRACTED BY EURACTIV BULGARIA

For the second release of the dataset, 135 case studies were based on the subject of climate change. The vast majority of them are from Bulgarian media, spanning from 2022 to 2024, with a strong concentration in 2023 and 2024. A small proportion are from English-language sources, exclusively on social media.

The sources are predominantly Bulgarian online media outlets and websites, with a smaller percentage coming from international posts on Twitter, Facebook, and YouTube videos.

It must be stressed that in the Bulgarian language, there are hundreds of so-called online "media", which are nothing else but instruments for disinformation, often, judging by their content, clearly to the service of Russia. In Bulgaria, the media regulator keeps their eyes wide closed on the phenomenon.

In Bulgaria, the Council of Electronic Media, the main regulator, has been taking the view for many years that "all opinions should be heard". As a result, charlatans have been given the floor on prime time, and even mainstream media ("Trud", "Epicenter", "Telegraph", "Struma") have quoted them, without a critical approach – normally they should have challenged their statements against views by reputed scientists.

Fake news in the Bulgarian language is much easier to find, especially in media outlets, compared to the Western context, where media are aware of the risk of fake news and make efforts to

become immune. In the Western context, social media platforms are the domain where fake news – including those relating to climate change – continues to prosper, especially on the X platform.

The primary focus of the fact-checks analysed is on criticizing the Green Deal, fostering anti-EU sentiment, and promoting the idea that anything related to environmental initiatives is unnecessary, costly, and harmful. They also seek to cultivate distrust in international organisations. In many instances, these narratives aim to discredit Western policies and values, undermining global efforts to address climate change and environmental sustainability. Some of the narratives draw on broader European issues, while others leverage local Bulgarian contexts relate to the war in Ukraine, to convey the same message - criticizing green initiatives and discrediting the European and Western agenda as a whole, and in addition – discouraging public opinion from the idea that Ukraine needs international help.

The fact-checks reveal a broader trend of misinformation and disinformation, particularly surrounding climate change, environmental policy, and socio-political issues. Several narratives, such as "Global Warming is a Lie" and "Russian Experts Predict a New Ice Age," contribute to climate change denial, a topic that continues to be polarising. These stories often position environmental activism and policies as economically harmful, with sensational claims like the "EU forcing veganism" or the "green mafia" destroying Bulgaria's economy. Such narratives aim to undermine public trust in environmental initiatives by exaggerating their negative impact.

Another notable theme is the rise of conspiracy theories. Claims about the HAARP project causing earthquakes, WEF's plan to slaughter pets to combat climate change, and even discussions of cannibalism on Swedish television reflect the exploitation of fear and confusion to distort legitimate environmental concerns. These narratives serve to distract from genuine climate crises and create a sense of helplessness or scepticism toward scientific efforts.

Additionally, there is an anti-European sentiment intertwined with many of these fact-checks, where EU policies are portrayed as economically oppressive or ideologically extreme, further fuelling public distrust in international organizations. The framing of economic and energy issues, particularly regarding the European energy crisis and the role of coal, positions environmental policies as a direct threat to national interests.

Overall, these fact-checks show a media landscape filled with alarmism, exaggeration, and conspiracy, which hinders informed public debate on climate and environmental issues. This spread of misinformation undermines scientific consensus and policy efforts, making it harder for the Bulgarian public to engage constructively with critical global challenges.

5   DISINFORMATION CASES COLLECTED IN THE THIRD RELEASE

In the third release of the datasets, 450 disinformation cases have been collected: 208 cases related to the war in Ukraine and 242 cases related to climate change.

In the following sub-sections, a brief description of the disinformation cases on the war in Ukraine and climate changes extracted by CNR, Euractiv Bulgaria, and Internews Ukraine are illustrated.

5.1   DISINFORMATION CASES ON THE WAR IN UKRAINE

5.1.1   DISINFORMATION CASES EXTRACTED BY INTERNEWS UKRAINE

Russian media have been systematically disseminating disinformation campaigns across online media and Telegram channels to undermine Ukraine during the ongoing war. This report analyzes key thematic groups of disinformation, outlining the narratives, frequency, and specific examples used to manipulate public opinion, discredit Ukraine, and destabilize support for the country.

The identified themes include: discrediting the Ukrainian government and Armed Forces, with frequent false claims about corruption, forced evacuations, and military failures; sowing mistrust towards Ukrainian refugees, portraying them as a burden or source of crimes abroad; and manipulating religious and cultural issues to incite societal divisions. Additionally, Russian propaganda amplifies accusations of Ukrainian "provocations" and "aggression," including alleged use of chemical weapons or plans to annex Russian territories (claims about Kursk region).

Other recurring narratives involve economic disinformation about Ukraine's financial crisis, attacks on Western partners by claiming reduced military support, and efforts to demoralize Ukraine's population through fabricated stories of protests, instability, or exaggerated casualty figures. The report highlights over 50 examples of these narratives, showcasing how Russian propaganda targets Ukraine's resilience, global partnerships, and societal unity. Each disinformation claim is analyzed and debunked using verified evidence and statements from reliable sources, ensuring a comprehensive understanding of these manipulative strategies.

1.   Mobilization and Military-Related Manipulations

A recurring narrative across Telegram channels and other social media platforms claims like "Don't let Zelensky mobilize himself."

Narratives about forced mobilization, fear, and military failures are consistently used to undermine trust in Ukrainian leadership and military institutions. The theme attempts to portray mobilization as indiscriminate and unfair, while suggesting demoralization within the armed forces.

2. Ukrainian Leadership and Government Criticism

Examples of narratives:

- The Ukrainian Parliament has lost its legitimacy.
- Zelensky has "acknowledged" the loss of Crimea.
- The Ukrainian leadership is restricting the president's access to information.

These narratives aim to delegitimize Ukrainian leadership, portraying the government as weak, compromised, or untrustworthy. Disinformation suggests internal conflicts and failures of governance.

3. Ukrainian Refugees and International Relations

Examples of narratives:

- Measles outbreak in Switzerland due to Ukrainian refugees.
- Germany worsens economy by supporting Ukraine

The theme focuses on discrediting Ukrainian refugees and portraying them as a burden or problem in host countries. It seeks to erode international support for Ukraine by exacerbating resentment among foreign populations.

4. Church and Religious Manipulations

Examples of narratives:

- The Spaso-Preobrazhensky Cathedral in Chernihiv was turned into a cinema.
- A UOC building in Lutsk was converted into a laundry.

Religious narratives aim to stoke societal division by exploiting tensions between the Ukrainian Orthodox Church (UOC) and the Orthodox Church of Ukraine (OCU).

5. Fake Atrocities and Human Rights Violations

Examples of narratives:

- Ukraine allegedly created "concentration camps" in the Sumy region.
- Ukraine forcibly takes blood from children for military needs.
- Ukrainian authorities restrict access to medical aid without updated military registration data.

These fabrications attempt to depict Ukraine as violating human rights, conducting unethical actions, and engaging in symbolic provocations against Russia or the West.

6. Economic and Social Instability in Ukraine

Examples of narratives:

- Ukrainian waters are polluted with corpse toxins due to growing cemeteries.
- Ukraine lacks any real "plan for victory."
- Ukrainian universities cut budget slots for conscription-age men.

These narratives emphasize economic and social collapse in Ukraine, seeking to discourage resistance and paint the government as incapable of governance during wartime.

7. False Allegations of Military Ineffectiveness

Examples of narratives:

- Actions by the Ukrainian army in Kursk were the most unjustified military operation of the 21st century.
- Supplies of F-16 jets or ATACMS missiles will not change the battlefield — U.S. Secretary of Defense.

These stories attempt to frame Ukraine's military strategy as futile and portray leadership as divided and ineffective.

The most frequent themes include military manipulations, leadership criticism, and refugee disinformation. These narratives target Ukraine's internal stability, international support, and military morale, often blending exaggerations, falsehoods, and emotional triggers to manipulate public perception. Internally, narratives aim to undermine national unity, pit citizens against their leadership, and erode trust in institutions. Internationally, the focus shifts toward discrediting Ukraine's government, discouraging foreign aid, and weakening solidarity among allies. By combining fabricated narratives with carefully targeted disinformation, propagandists seek to weaken Ukraine's internal cohesion while eroding international confidence and support. Such efforts underscore the importance of vigilance, media literacy, and coordinated counter-narratives to combat these destabilizing tactics.

5.1.2 DISINFORMATION CASES EXTRACTED BY EURACTIV BULGARIA

For the third release of the dataset, Euractiv Bulgaria analysed and dissected according to the agreed template a total of 63 claims on the war in Ukraine.

The war in Ukraine is for the Bulgarian media a much bigger topic than climate change, although both issues are treated with the same bias in all the websites mentioned above.

Many of the articles analysed have been translated from Russian or adapted from Russian sources.

Overall, the articles analysed are aimed at creating panic that NATO or the US is sending Bulgaria at war with Russia, that Ukraine is a corrupt country that doesn't deserve any help, sometimes the government in Kyiv is named "Nazi regime" just as in Russian media. Western experts are quoted (without evidence that they made such unsane statements) that it is high time NATO attacks Russia.

The Crocus City Hall terrorist attack in Moscow on 22 March 2024 has prompted many publications suggesting that it was the West and Ukraine, not the Islamist terrorists, that were behind it.

Many of the articles contain insulting qualifications with regard to Zelenskyy, who is ridiculed "by default".

Some of the messages conveyed by the claims analysed related to the Ukraine war are as follows:

- Many of the strikes against Russia started from Bulgarian territory (US planes located in Bulgarian airfields allegedly striking Russian targets) which makes Bulgaria a legitimate target;

- French troops have already arrived in Ukraine, coming from Bulgarian territory;

- Bulgaria gave precious military equipment to Ukraine and is now depleted of defence capabilities, Bulgaria sends such equipment without the necessary vote in parliament;

- The reformist Bulgarian government of Kiril Petkov was preparing new taxes to finance the support to Ukraine;

- Criminal groups from Ukraine penetrate Bulgaria;

- Zelenskyy tramples Orthodoxy to the extent that he is presented as the anti-Christ;

- Zelenskyy is corrupt (yachts) and so is his wife (Bugatti); Zelenskaya is involved in trafficking of Ukrainian children for paedophiles in Europe;

- The mental health of Zelenskyy is questioned;

- The protests of farmers across Europe are presented as a reaction to the impact they suffer from Ukraine imports;

- Ukraine has become a haven or a holiday destination for terrorists;

- Ukraine's rulers won't last, a coup is imminent; Zelenskyy is ready to flee, he has transferred his billions to Cyprus, Switzerland and Albania;

- Kyiv ordered the attempted assassination of Donald Trump in Pennsylvania.

5.2 DISINFORMATION CASES ON CLIMATE CHANGE

5.2.1 DISINFORMATION CASES EXTRACTED BY CNR

In the latest months, climate change disinformation has further intensified, posing a significant challenge to public understanding and policy development. In fact, recent reports highlight how disinformation has evolved from outright denial of climate change to undermining proposed solutions, such as renewable energy initiatives, or promoting events that seeks to stall policy action. This shift often manifests as campaigns (even political ones) that discredit renewable technologies or exaggerate the costs and inefficiencies of green policies. Social media platforms remain major conduits for spreading such narratives.

In the attempt to face these issues, CNR has collected further 150 claims related to climate change misinformation for the third release of the database.

All the analysed claims were extracted from certified and recognized fact-checking websites. The websites have been mainly selected from the repository of the European Digital Media Observatory (EDMO) and eufactcheck.eu.

In particular, the selected fact checking websites were:

- eufactcheck.eu
- factcheck.vlaanderen
- correctiv
- facta.news
- afp factuel

As the nature of this specific topic requires specialized feedback coming from the global scientific community, several claims have also been extracted from the following fact-checking websites:

- skepticalscience
- the daily sceptic
- Politifact.com
- Kritichno.bg
- science.feedback.org

The main collected topics were concerned with the following misleading narratives:

- **Questioning Scientific Consensus**: Misinformation suggesting that scientists are divided on the causes or reality of climate change, despite overwhelming evidence to the contrary. This is often used to undermine trust in climate science.
- **Blaming Natural Cycles**: Narratives asserting that extreme weather events or global temperature changes are due to natural variability rather than human activity.
- **"Greenwashing"**: Corporate campaigns that exaggerate the environmental benefits of their products or operations, distracting from their contributions to emissions and environmental degradation.
- **Attacks on Renewable Energy Projects:** Claims that renewable energy projects harm ecosystems more than they benefit them, often overstating the negative impacts while ignoring long-term sustainability advantages.
- **Economic Alarmism**: Claims that climate policies will harm the economy, lead to mass job losses, or drastically increase energy costs. These arguments often overlook the economic benefits of renewable energy and the long-term costs of inaction.
- **Misleading narratives coming from political sources:** These narratives about often aim to protect vested interests or undermine policy changes. Also, some politicians attribute climate changes to natural cycles rather than human activity, disregarding scientific consensus.

5.2.2   DISINFORMATION CASES EXTRACTED BY EURACTIV BULGARIA

For the third release of the dataset, Euractiv Bulgaria analysed and dissected according to the agreed template a total of 93 claims on climate change.

As previously said, fake news in Bulgaria often come from the country's officials. As an example, Bulgaria has had an environment minister, Neno Dimov (2017-2020) who makes no secret that he is a climate change denier. He continues to make controversial statements, as our analysis shows.

In a more recent period, the political party Velichie (Greatness) gained prominence and acceded to Parliament in June 2024 thanks to mushroom websites (the mother of which is Krasivovetrino.bg) spreading messages against wind farms, against the Green Deal and EU policies in general. Several of our claims analyse the activity of these websites. The mushroom websites attack the pro-Western politicians Kiril Petkov, Assen Vassilev, and Daniel Laurer for having a vested interest in wind turbine projects and says they will "push them through by force".

An ever-fresher political presence is that on Kuzman Iliev, a self-proclaimed expert in energy and climate policies, who often argues against the EU Green Deal and against Bulgaria's accession to the eurozone. Iliev, who is a protagonist in our research, just founded a political party called Bulgaria mozhe (Bulgaria can).

Two scientists from the Bulgarian Academy of Sciences, Katia Georgieva and Boyan Kirov, often make statements denying climate change, using a semblance of scientific arguments. But it's not their arguments, it's mostly the fact that they represent an official institution that makes their statements credible for many.

Another trick is to present climate-related issues from the side of ecologists, while doing exactly the opposite. The website Ekonovini.bg (EcoNews) says it approaches the topics related to ecology objectively. In reality, the website defends construction building in protected areas and presents environmental protection as harmful to business and the well-being of Bulgarians.

A source of fake news and disinformation regarding both the war in Ukraine and climate change are podcasts on YouTube, which the Euractiv Bulgaria team examined. On YouTube, under various names, the pro-Russian sect ALLAT Ra, banned in Ukraine, is presented. The sect also operates in Bulgaria and spreads claims that climate change has occurred and continues to occur independently of human activity.

Some of the messages conveyed by the claims analysed related to climate change are as follows:

- Climate change is a hoax, some are turning it into a religion, but in fact it's a muti-billion business, Bill Gates is the main profiteer, Russia and China will be the main winners, because they don't buy the narrative;

- the EU is committing suicide with the Green Deal, it is the reason why the price of energy in Europe is skyrocketing, by implementing the Green Deal we all will be soon eating insects, food and electricity will be rationed;

- Bulgarian politicians are yes-men, the closure of coal centrals as Brussels wants will be a catastrophe for the country;

- High electricity prices in Bulgaria are due to the country's electricity exports to Ukraine;

- Solar energy, and renewable energies in general, are polluting more than coal, wind energy will destroy landscapes and tourism;

- Home farms are targeted as harmful for the environment, the state will confiscate private ownership of land;

- Forget about climate policies, the end of global warming is coming (or global warming has already ended), Russian scientists say.

6 ASSESSMENT OF THE DATA QUALITY OF THE DATASETS

Both datasets have been assessed considering the Data Quality Assessment Framework (DQAF) (Sebastian-Coleman, 2012).

This framework relies on a set of prerequisites and five dimensions of data quality that are listed below:

1) assurances of **integrity,** which assesses the completeness and trustworthiness of the data;
2) **methodological soundness,** which examines whether the database is coherent with rigorous and consistent analytical methodologies;
3) **correctness and dependability**, which evaluates the consistency and accuracy of data retrieval and performance over time.
4) **serviceability,** which measures the database's responsiveness to user needs, including its adaptability and support infrastructure;
5) **accessibility,** which ensures that the database is user-friendly and readily available to authorized users.

Each dimension also comprehends a set of elements of good practice and several related indicators that allow to measure the level of accomplishment of the elements. In this direction, a thorough evaluation based on the selected criteria can provide a comprehensive understanding of the database's strengths and limitations related to the quality of the collected data.

For each abovementioned dimension, we reported the elements of good practice and the indicator specific for evaluating our two datasets, as shown in Table 2. The scores of the indicators range from *O = Practice Observed*, meaning that the indicator obtains a value equal or higher than 80%; *LO = Practice Largely Observed*, meaning that the indicator obtains a value from 50% and 80%; *LNO =Practice Largely Not Observed*, meaning that the indicator obtains a value from 20% to 50%; *NO = Practice Not Observed*, meaning that the indicator obtains a value lower than 20%; *NA = Not Applicable*.

Table 2. DQAF defined for evaluating our datasets

DIMENSION OF DATA QUALITY	ELEMENTS	INDICATORS	SCORES
1) ASSURANCES OF INTEGRITY: Upholding objectivity and	**1.1) PROFESSIONALISM**: Decisions guided by statistical principles; impartial data production.	**1.1.1) Statistics are produced on an impartial basis.** **1.2.1) Terms and conditions for data production are public.**	**O: Practice observed**: >= 80% **LO: Practice largely observed:** >= 50%; < 80%

ethical standards in collecting, processing, and disseminating data	**1.2) TRANSPARENCY**: Clear public guidelines, advance notices of changes, and identification of statistical products. **1.3) ETHICAL STANDARDS**: policies and practices are guided by ethical standards	**1.3.1) Guidelines for staff behavior are in place and are well known to the staff.**	**LNO Practice Largely Not Observed:** >= 20%; < 50% **NO: Practice Not Observed:** < 20%
2) METHODOLO GICAL SOUNDNESS: Adherence to internationally accepted standards, guidelines, or practices.	**2.1) CONCEPTS AND DEFINITIONS**: use of globally recognized statistical frameworks. **2.2) SCOPE, CLASSIFICATION, RECORDING**: consistency with accepted practices for sectors, flows, and stocks	**2.1.1) The overall structure in terms of concepts and definitions follows internationally accepted standards, guidelines, or good practices.** **2.2.1) the scope, classification, and recording are broadly consistent with internationally accepted standards, guidelines, or good practices.**	**O: Practice observed**: >= 80% **LO: Practice largely observed:** >= 50%; < 80% **LNO Practice Largely Not Observed:** >= 20%; < 50% **NO: Practice Not Observed:** < 20%
3) ACCURACY AND RELIABILITY: Ensuring outputs are accurate and reflect reality	**3.1) SOURCE DATA:** comprehensive, timely, and well-validated source data **3.2) STATISTICAL TECHNIQUES:** use of sound processing methods **3.3) ASSESSMENT AND VALIDATION OF INTERMEDIATE DATA AND STATISTICAL OUTPUTS**: intermediate results and statistical outputs are regularly assessed and validated.	**3.1.1) source data are obtained from comprehensive data collection programs that take into account country-specific conditions.** **3.2.1) data compilation employs sound statistical techniques to deal with data sources.** **3.3.1) Intermediate results are validated against other information where applicable.**	**O: Practice observed**: >= 80% **LO: Practice largely observed:** >= 50%; < 80% **LNO Practice Largely Not Observed:** >= 20%; < 50% **NO: Practice Not Observed:** < 20%

	3.4) REVISION STUDIES*: revisions, as a gauge of reliability, are tracked and mined for the information they may provide.	**3.4.1) Studies and analyses of revisions are carried out routinely**	
4) SERVICEABILITY: Data is useful, consistent, and available on time.	**4.1) PERIODICITY AND TIMELINESS:** data follows regular release schedules aligned with international standards **4.2) CONSISTENCY:** data remains consistent across time, datasets, and frameworks **4.3) REVISION POLICY:** clear procedures for updates and making revision studies public.	**4.1.1) periodicity and timeliness follows dissemination standards.** **4.2.1) statistics are consistent or reconcilable over a reasonable period of time.** **4.3.1) revisions follow a regular and transparent schedule AND preliminary and/or revised data are clearly identified.**	**O: Practice observed**: >= 80% **LO: Practice largely observed:** >= 50%; < 80% **LNO Practice Largely Not Observed:** >= 20%; < 50% **NO: Practice Not Observed:** < 20%
5) ACCESSIBILITY: Data and metadata are easily available to users	**5.1) DATA ACCESSIBILITY**: statistics presented clearly and disseminated impartially **5.2) METADATA ACCESSIBILITY:** up-to-date and pertinent metadata are made available. **5.2) USER SUPPORT**: prompt and knowledgeable assistance for users	**5.1.1) statistics are presented in a way that facilitates proper interpretation and meaningful comparisons.** **5.2.1) documentation on concepts, scope, classifications, basis of recording, data sources, and statistical techniques is available, and differences from internationally accepted standards, guidelines, or good practices are annotated.**	**O: Practice observed**: Value of the Indicator >= 80% **LO: Practice largely observed:** 50% <= Value of the Indicator < 80% **LNO Practice Largely Not Observed:** 20% <= Value of the Indicator < 50% **NO: Practice Not Observed:**

		5.3.1) catalogs of publications, documents, and other services, including information on any changes, are widely available.	Value of the Indicator < 20%

After defining the DQAF for assessing the data quality of our datasets, we evaluated the scores of the indicators listed in the third column of Table 2. The scores and motivation for each indicator, resulted from the assessment, as well as the total score of the dimensions are shown in Table 3.

Table 3. Scores of the DQAF indicators and motivations resulted from the assessment

INDICATORS	SCORE OF THE INDICATOR	MOTIVATION	TOTAL SCORE OF THE DIMENSION
1.1.1) Statistics are produced on an impartial basis	100%	Data have been collected, processed, and disseminated by avoiding external pressures. In order to maintain objectivity and avoid bias, the collection and processing has been conducted independently by different professional experts. Indeed, collaboration among independent experts helped maintain an unbiased and objective dataset.	
1.2.1) Terms and conditions for data production are public	100%	The procedural framework governing data collection and release is accessible to users. The deliverable 6.1, containing the procedural framework followed for the data production and collection, is available open access. This ensures transparency, allowing stakeholders to review and verify research methods and findings.	O: Practice observed
1.3.1) Guidelines for staff behavior are in place and are	50%	A template has been given to professional experts conducting the collection of data. However, no guidelines for staff behaviour have been defined and distributed.	

well known to the staff			
2.1.1) The overall structure in terms of concepts and definitions follows internationally accepted standards, guidelines, or good practices	100%	The data concepts and definitions follow guidelines and established global standards. In particular, the theoretical framework defined in WP4 has been followed, which relies on a comprehensive analysis of accepted standards and good practices.	
2.2.1) The scope, classification, and recording are broadly consistent with internationally accepted standards, guidelines, or good practices	100%	The scope, classification and recording of data align with international frameworks. In particular, for the scope it has been clearly defined what constitutes "fake news" within the dataset and specified the thematic focus (e.g., disinformation about the war in Ukraine, and climate change). For the classification, a standard taxonomy has been used that categorizes fake news by topic, source, keyword, etc. (see Table 1). Finally, for the recording of data we used standard file formats (csv and exe for the overall datasets, jpeg and png for images, mp4 for videos) and metadata to capture the date of publication, the original language and any translations, the URLs, and the fact-checking status.	**O: Practice observed**:
3.1.1) Source data are obtained from comprehensive data collection programs that take into account country-specific conditions	100%	Data sources are robust and representative of the intended topic and take into account country-specific conditions. Indeed, the majority of data is collected from highly reputable fact checking websites (e.g., EDMO) and verified news websites and social media fact-checked by professional journalists.	**LO: Practice largely observed**

3.2.1) Data compilation employs sound statistical techniques to deal with data sources	0%	No statistical methods are employed for data collection, estimation, and analysis of the two datasets. Such techniques will be used in Task 6.2 for enriching the two datasets by extracting further relevant features through Natural language processing, sentiment analysis, topic modelling, etc.	
3.3.1) Intermediate results are validated against other information where applicable	100%	Two intermediate results have been provided (dataset releases at month 6 and 9). These intermediate results have not been validated by cross-referencing them with external or auxiliary sources of information, because they have been collected from reliable fact-checking organizations or fact-checked by professional journalists. Therefore, the accuracy and reliability of intermediate results is guaranteed.	
3.4.1) Studies and analyses of revisions are carried out routinely	0%	No updates to fact-checking outcomes (e.g., new evidence refutes or verifies previously flagged content) are foreseen in the project.	
4.1.1) Periodicity and timeliness follow dissemination standards	100%	The dataset adheres to pre-established international dissemination schedules in accordance with pre-announced dissemination timelines. In fact, the datasets are released and disseminated following the timeline defined in the Grant Agreement (3 releases are delivered). These procedures enable consistent reporting and cross-border cooperation by demonstrating a commitment to international cooperation, timeliness and transparency.	**LO: Practice largely observed**
4.2.1) Statistics are consistent or reconcilable over a reasonable period of time	0%	No standardized methodologies for defining and identifying fake news across different time periods, no documentation of changes, and no trend monitoring have been applied to ensure that data remains reliable, comparable, and interpretable over time.	

4.3.1) Revisions follow a regular and transparent schedule and preliminary and/or revised data are clearly identified.	100%	Updates and revisions of the datasets are made according to an established and publicly accessible schedule. Moreover, datasets are clearly marked to distinguish preliminary from finalized data. This ensures that all changes are thoroughly recorded, allowing users to monitor updates and keep a clear overview of how the datasets have changed over time.	
5.1.1) Statistics are presented in a way that facilitates proper interpretation and meaningful comparisons.	100%	Deliverable 6.1 contains data tables, charts, and figures that are presented in a clear format that facilitates interpretation, even if no comparisons across sectors and countries are performed.	
5.2.1) documentation on concepts, scope, classifications, basis of Recording, data sources, and statistical techniques is available, and differences from internationally accepted standards, guidelines, or good practices are annotated.	50%	A clear and comprehensive documentation on concepts, scope, classifications, basis of recording, data sources is available within the Deliverables 4.1 and 6.1. However, documentation on statistical techniques is not available, and differences from internationally accepted standards, guidelines, or good practices have not been annotated.	**O: Practice observed**
5.3.1) Catalogs of publications, documents, and other services,	100%	Relevant resources, datasets, and updates are easily available online or through other media. Indeed, the two datasets will be freely available and easy to download with	

including information on any changes, are widely available.	detailed documentation on Zenodo and GitHub.

The results of the data quality assessment we performed on the two datasets are summarized in Table 4.

Table 4. Data Quality Assessment Framework - Summary Results		
Key to symbols: O = Practice Observed; LO = Practice Largely Observed; LNO =Practice Largely Not Observed; NO = Practice Not Observed; NA = Not Applicable;		
Datasets **Dimensions/Elements**	**Ukrainian War dataset**	**Climate change dataset**
1. Integrity		
1.1 Professionalism	**O**	**O**
1.2 Transparency	**O**	**O**
1.3 Ethical standards	**LO**	**LO**
2. Methodological soundness		
2.1 Concepts and Definitions	**O**	**O**
2.2 Scope, Classification, recording	**O**	**O**
3. Reliability		
3.1 Source data	**O**	**O**
3.2 Statistical Techniques	**NO**	**NO**
3.3 Assessment and validation of intermediate data and statistical outputs	**O**	**O**
3.4 Revision studies	**NO**	**NO**
4. Serviceability		
4.1 Periodicity and Timeliness	**O**	**O**

4.1 Consistency	**NO**	**NO**
4.2 Revision policy	**O**	**O**
5. Accessibility		
5.1 Data accessibility	**O**	**O**
5.2 Metadata accessibility	**LO**	**LO**
5.3 User support	**O**	**O**
Practice observed: Current practices generally meet or achieve the objectives of DQAF internationally accepted statistical practices without any significant deficiencies. Practice largely observed: Some departures, but these are not seen as sufficient to raise doubts about the authorities' ability to observe the DQAF practices. Practice largely not observed: Significant departures and the authorities will need to take significant action to achieve observance. Practice not observed: Most DQAF practices are not met. Not applicable: Used only exceptionally when statistical practices do not apply to the country's circumstances.		

7   CONCLUSIONS

This deliverable introduced the datasets prepared for collecting fake statements and related multimedia contents on the two case studies considered within the project (i.e., the war in Ukraine and the climate changes).

The structure and implementation of the two datasets were introduced by describing (i) the template used for extracting data from disinformation cases, (ii) the Database Management System (DBMS) used for the implementation of the datasets, and (iii) the cloud environment used for storing and sharing the multimedia contents. Finally, a brief description of disinformation cases collected and analysed for the three releases of the datasets was provided.

The first release of these datasets contains 1000 disinformation cases, of which 500 cases are related to the war in Ukraine and 500 cases related to climate change. The main disinformation misleading narratives about the war in Ukraine ranged from the third world war to a biological disaster, targeting the Ukrainian army and President Volodymyr Zelensky, questioning the authenticity of the conflict, disinformation about refugees, territorial integrity, and sovereignty, military and security threats, and human rights violations and internal governance. Regarding climate changes, the main collected topics were concerned with misleading narratives about solar activity and climate changes, CO_2's role in climate changes, global warming and ice melting, electric vehicles, agriculture, climate, and the green deal, Greta Thunberg's activity against climate changes, volcano activity and climate changes, chemical trails, geoengineering, and electric vehicles.

The second release of these datasets contains 550 disinformation cases, of which 215 cases related to the war in Ukraine and 335 cases related to the climate change. The main disinformation misleading narratives about the war in Ukraine ranged from the counter-offensive related disinformation, discrediting the Ukrainian government, compromising Ukrainians' image, legitimizing Russia and the invasion, attacking NATO and Ukraine foreign supporters, and eroding public trust in international alliances. Regarding climate change, the main collected topics were concerned with wind turbines, conspiracy theories, misrepresentation of data, scepticism of scientific consensus, denial of climate change, discrediting renewable energy, misleading narratives about policy impacts, criticizing the Green Deal and other European green initiatives, undermining European environmental policy efforts.

In the third release of the datasets contains 450 disinformation cases, of which 208 cases related to the war in Ukraine and 242 cases related to climate change. The main disinformation misleading narratives about the war in Ukraine ranged from discrediting the Ukrainian government and Armed Forces to sowing mistrust towards Ukrainian refugees, manipulating religious and cultural issues to incite societal divisions, economic disinformation about Ukraine's

financial crisis, attacks on Western partners by claiming reduced military support, efforts to demoralize Ukraine's population, creating panic that NATO or the US is sending Bulgaria at war with Russia, and discreditation of Ukraine seen as a corrupt country that doesn't deserve any help. Regarding climate change, the main collected topics were concerned with questioning scientific consensus, blaming natural cycles, "Greenwashing, attacks on renewable energy projects, economic alarmism, misleading narratives coming from political sources, climate change is a hoax, the EU is committing suicide with the Green Deal, and Bulgarian politicians are yes-men.

REFERENCES

Zhang X, Ghorbani AA. (2019). An overview of online fake news: characterization, detection, and discussion. *Information Processing & Management* 57(2):102025

D'Ulizia, A., Caschera, M. C., Ferri, F., & Grifoni, P. (2021). Fake news detection: a survey of evaluation datasets. *PeerJ Computer Science*, *7*, e518.

Taipalus, T. (2023). Database management system performance comparisons: A systematic literature review. *Journal of Systems and Software*, 111872.

Sebastian-Coleman, L. (2012). Measuring data quality for ongoing improvement: a data quality assessment framework. Newnes.

Review Sheet of Deliverable/ Milestone Report

D6.1: Starting dataset of fake statements and related multimedia contents

Editor(s):	CNR, IUA, Euractiv
Responsible Partner:	CNR
Status-Version:	Draft
Date:	02/01/2025
Distribution level (CO, PU):	PU
Reviewer (Name/Organization)	Pascaline Gaborit, Pilot4dev
Review date	

Disclaimer: This assessment reflects only the author's views and the European Commission is not responsible for any use that may be made of the information contained therein"

Mark with X the corresponding column:

Y= yes	N= no	N = not applicable

ELEMENT TO REVIEW	Y	N	NA	COMMENTS
FORMAT: Does the document … ?				
…include editors, deliverable name, version number, dissemination level, date, and status?				
…contain a license (in case of public deliverables)?				
…include the names of contributors and reviewers?				
….has a version table consistent with the document's revision?				
… contain an updated table of contents?				
… contain a list of figures consistent with the document's content?				
… contain a list of tables consistent with the document's content?				
… contain a list of terms and abbreviations?				
… contain an Executive Summary?				
… contain a Conclusions section?				
… contain a List of References (Bibliography) in the adequate format, if relevant?				
… use the fonts and sections defined in the official template?				
… use correct spelling and grammar?				
… conform to length guidelines (50 pages maximum (plus Executive Summary and annexes)				
… conform to guidelines regarding Annexes (inclusion of complementary information)				
… present consistency along the whole document in terms of English quality/style? (to avoid accidental usage of copy&paste text)				
About the content…				
Is the deliverable content correctly written?				
Is the overall style of the deliverable correctly organized and presented in a logical order?				
Is the Executive Summary self-contained, following the guidelines and does it include the main conclusions of the document?				
Is the body of the deliverable (technique, methodology results, discussion) well enough explained?				
Are the contents of the document treated with the required depth?				
Does the document need additional sections to be considered complete?				
Are there any sections in the document that should be removed?				

ELEMENT TO REVIEW	Y	N	NA	COMMENTS
Are all references in the document included in the references list?				
Have you noticed any text in the document not well referenced? (copy and paste of text/picture without including the reference in the reference list)				
SOCIAL and TECHNICAL RESEARCH WPs (WP4, 5, 12, 13, 14)				
Is the deliverable sufficiently innovative?				
Does the document present technical soundness and its methods are correctly explained?				
What do you think is the strongest aspect of the deliverable?				
What do you think is the weakest aspect of the deliverable?				
Please perform a brief evaluation and/or validation of the results, if applicable.				
AI AND TECNOLOGICAL WPS (WP6 – WP11)				
Does the document present technical soundness and the methods are correctly explained?				
What do you think is the strongest aspect of the deliverable?				
What do you think is the weakest aspect of the deliverable?				
Please perform a brief evaluation and/or validation of the results, if applicable.				
DISSEMINATION AND EXPLOITATION WPs (WP15 – WP17)				
Does the document present a consistent outreach and exploitation strategy?				
Are the methods and means correctly explained?				
What do you think is the strongest aspect of the deliverable?				
What do you think is the weakest aspect of the deliverable?				
Please perform a brief evaluation and/or validation of the results, if applicable.				

SUGGESTED IMPROVEMENTS

PAGE	SECTION	SUGGESTED IMPROVEMENT
		.

CONCLUSION
Mark with X the corresponding line.

	Document accepted, no changes required.
	Document accepted, changes required.
	Document not accepted, it must be reviewed after changes are implemented.

Please rank this document globally on a scale of 1-5 (1 = poor, 5= excellent) – using a half point scale. Mark with X the corresponding grade.

Document grade	1	1.5	2	2.5	3	3.5	4	4.5	5